Social Effects by the Singularity

～ Pre-Singularity Era ～

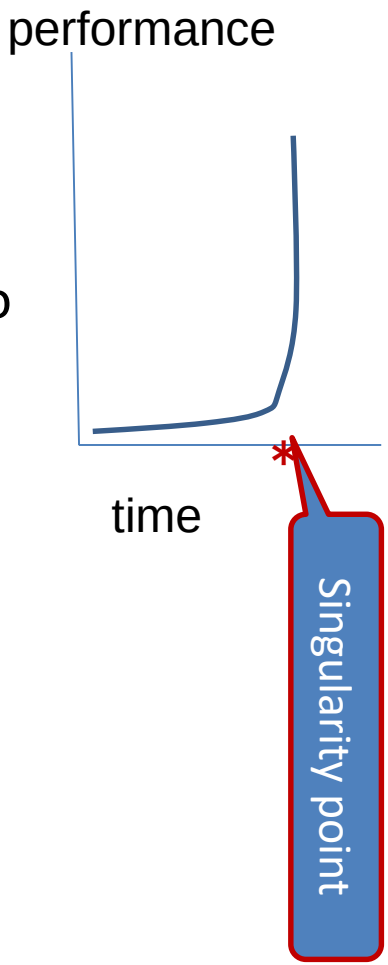
Hiroshi Nakagawa

Reference

- Ray Kurzweil: The Singularity is Near , Loretta Barrett Books Inc.2005
- James Barrat : Our Final Invention: Artificial Intelligence and the End of the Human Era, *New York Journal of Books*, 2013
- Nick Bostrom: Superintelligence, Oxford University Press. 2014
- John Markoff: Machines of Loving Grace: The Quest for Common Ground Between Humans and Robots ,2015
- Thomas H. Davenport , Julia Kirby :Only Humans Need Apply: Winners and Losers in the Age of Smart Machines , Harper Business, 2016

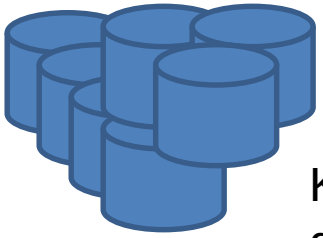
What is the Post Human that comes after the singularity?

- Kurzweil writes that the singularity is :
  - Technologies about G（gene）、N（nanotech.）、R（robot including AI）develop exponentially as depicted in the right figure
  - ＊ is singularity point. After it, tech. grows rapidly to the infinity



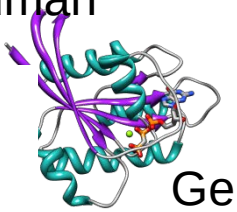
- Kurzweil predicts the singularity will occur in 2045.
- Post human is a new version of human that is transformed by G, N, and R after singularity point. → Not a human anymore

Image of post human

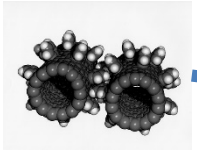


Knowledge is up or down loaded between the brain and outside via internet.

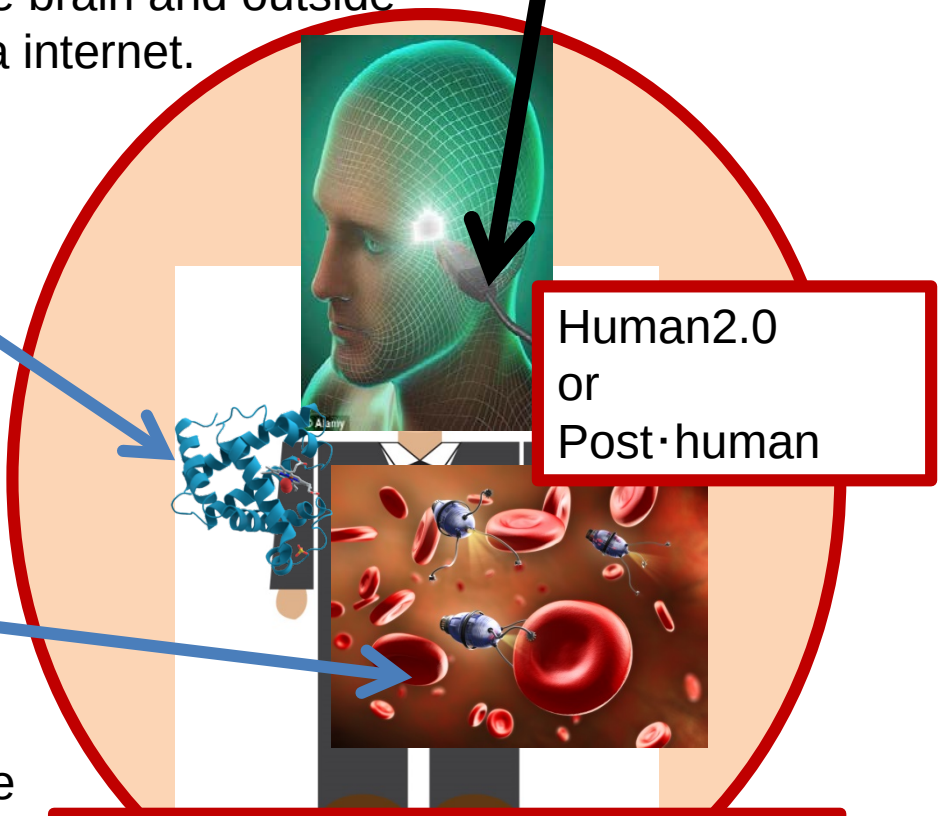
Human 1.0
Today's human



Gene transformation makes completely immune from disease, and even immortal.



Nano-bots made from carbon nano tube are injected into blood stream, heal the disease parts and communicate to outside of body.



Human2.0
or
Post·human

Immortal and acquire all the knowledge all people have.

Pre-Singularity

- Is a post human really possible?
  - Puerto Rico Conference on 2014 organized by computer scientists said it is impossible.
    - This conference is a counter part of Asilomar Conference of bio-scientists.
  - Especially, nano-bot is so hard to realize that Super Artificial Intelligence(Super AI) which is much more intelligent is necessary to develop nano-bot.
  - Some researchers say that apoptosis should be implanted to AI in order for Super AI not have proliferation（Barrat Ch. 14)
    - Apoptosis is a process of programmed cell death that occurs in multicellular organisms.

Pre-Singularity

- Even in Pre-Singularity era, meaning the time before the singularity point, development of advanced AI will causes various types of social, ethical , technological, or even humanity related problems.

- The remaining part shows these problems and hopefully the solution to these.
  - Unfortunately, hope is very small.

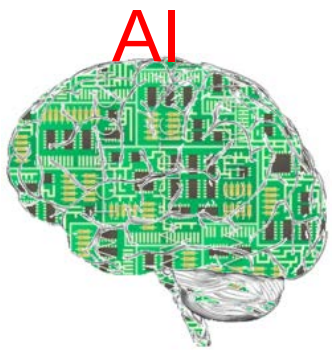
Contents

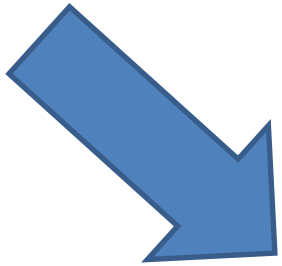
- Stance of scientists community against Pre-Singularity problems
- Amplification vs. Replacement
- AI takes over jobs
- Boarder line between amplification and replacement
  - Autonomous driver : trolley problem
  - The right to be forgotten
- Towards black box
  - Responsibility
  - Vulnerability of financial dealing system made of many AI agent traders connected via internet
- AI and weapon
- Filter bubble phenomena
- Analogy : Selfish gene
- AI and privacy
  - The right to be forgotten, Profiling and Don't Track
  - Feeling of friendliness to android
- Again self conscious and identity

Stance of scientists community

- Asilomar Conference held by computer scientists in 2009 rejected the Kurzweil's idea of self autonomous AI realizes.（Markoff Ch.10）
- Google had acquired Deep Mind in 2014, then computer scientists organized Puerto Rico Conference on 2014, but they were extremely passive compared to Asilomar of Bio-science.（Markoff Ch.10）
- Many people think it problem that computer scientist's community does not have a sense of crisis about AI.

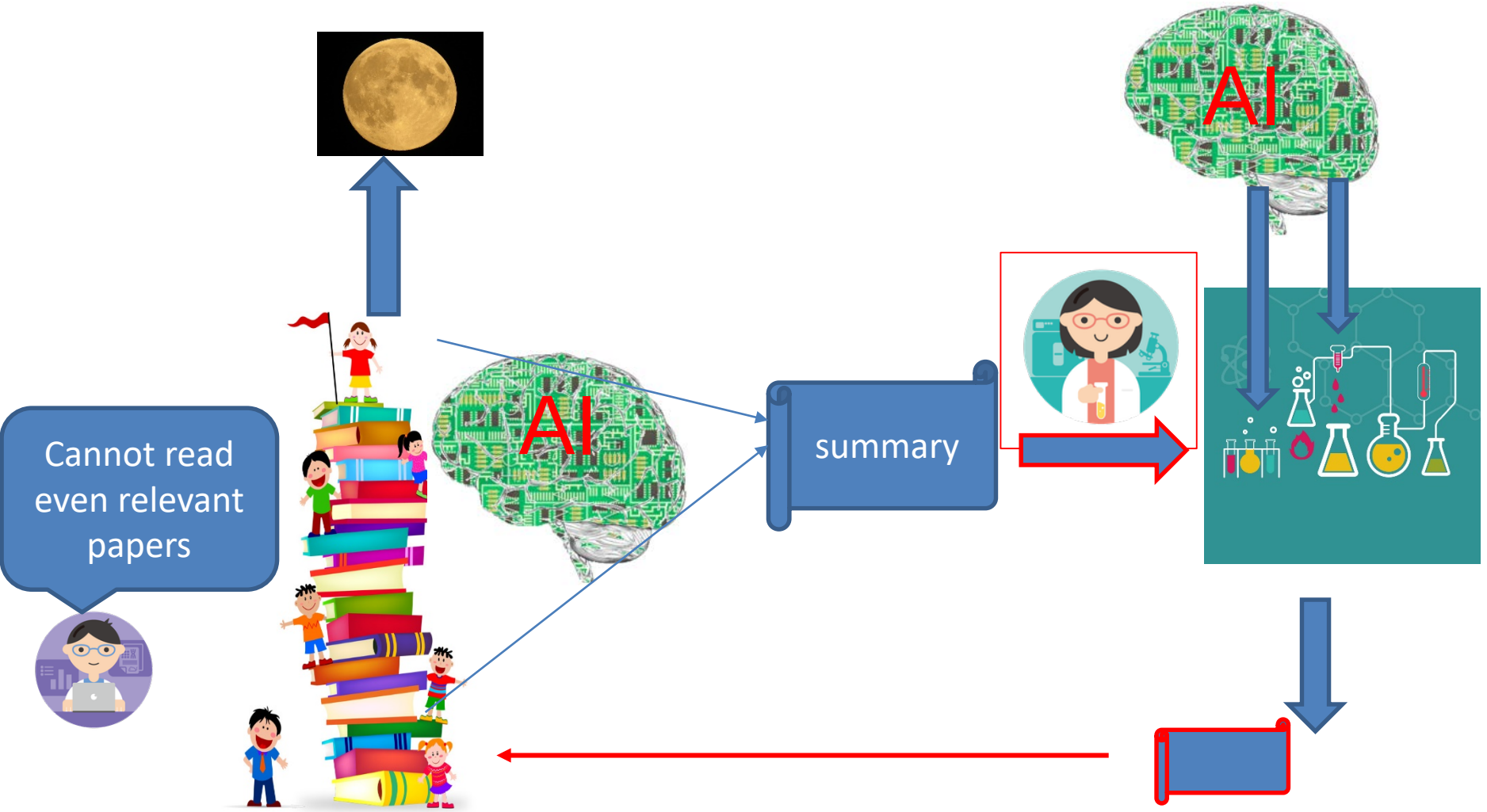
AI takes over human jobs





Where, What,
How???

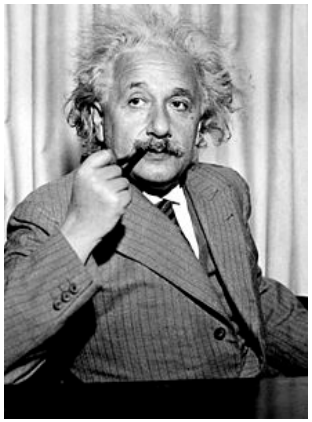
Researching job that are thought to be most distant from this crisis are seemingly taken over by AI.



Davenport says we can find new jobs immune to AI invasion, but…

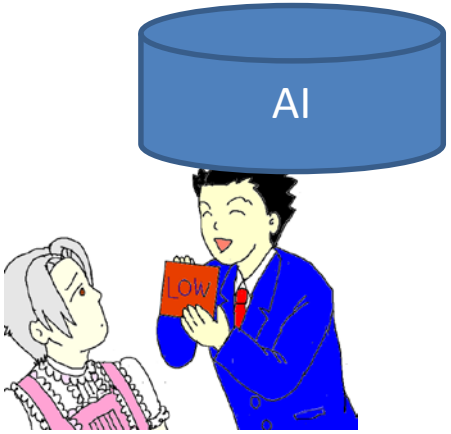
- Jobs aiming at higher quality than AI, say judgment without data

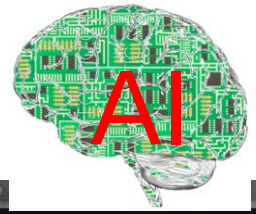
➢ It must be done by so called genius?

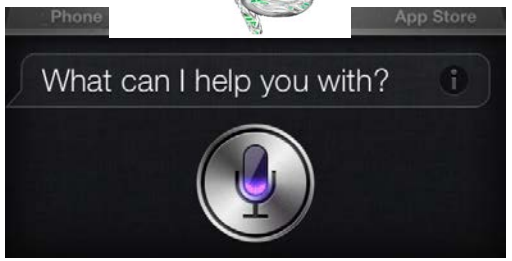


Davenport says we can find new jobs immune to AI invasion, but…

- Jobs AI can not do such as human intercommunication, persuasion, etc..
  - AI knows more and precise about things, event, etc.
  - If NLP becomes great , AI is more relied by human…



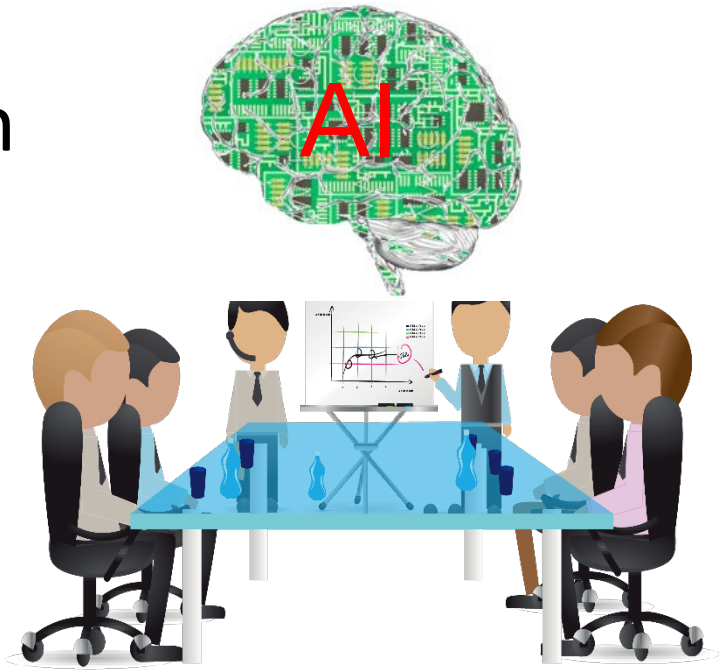




Davenport says we can find new jobs immune to AI invasion, but…

- Jobs of finding connection between business and technology

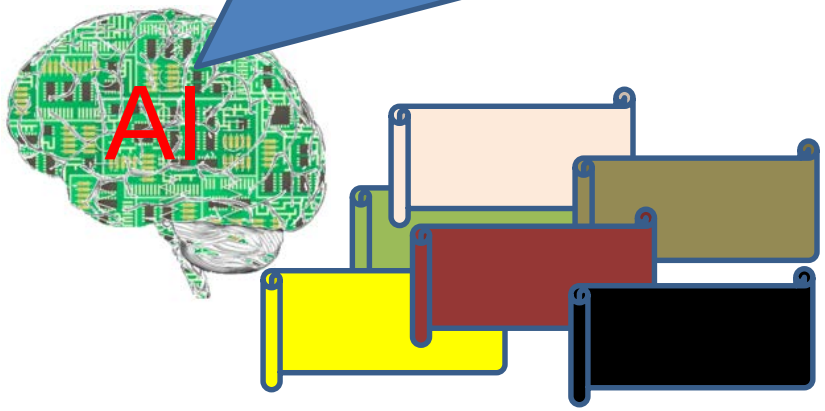
- AI can do more thorough investigation about linkage between business and technology than human.



Davenport says we can find new jobs immune to AI invasion, but…

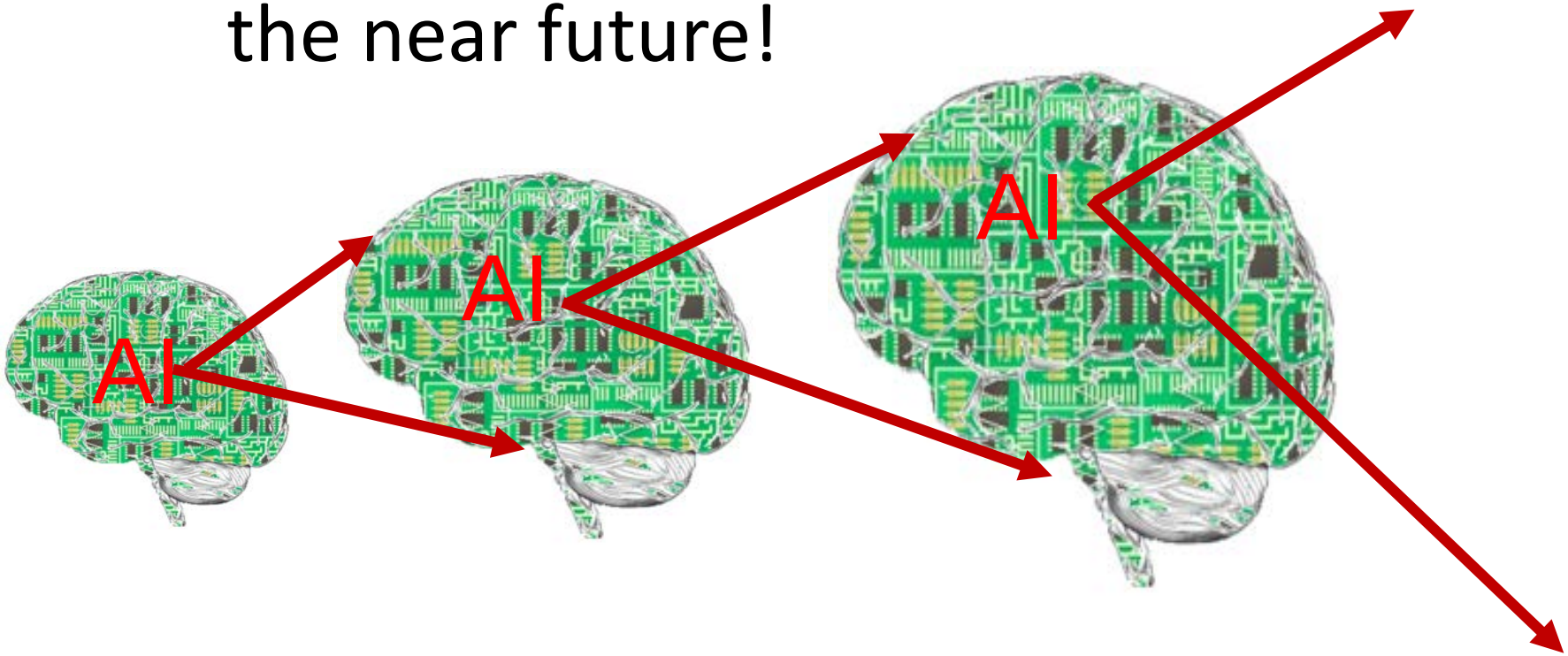
- Jobs less economical with employing machine or AI , in other words, quite rare and case specific task.

AI has a learning tech, so easy to cope with rare or special task!



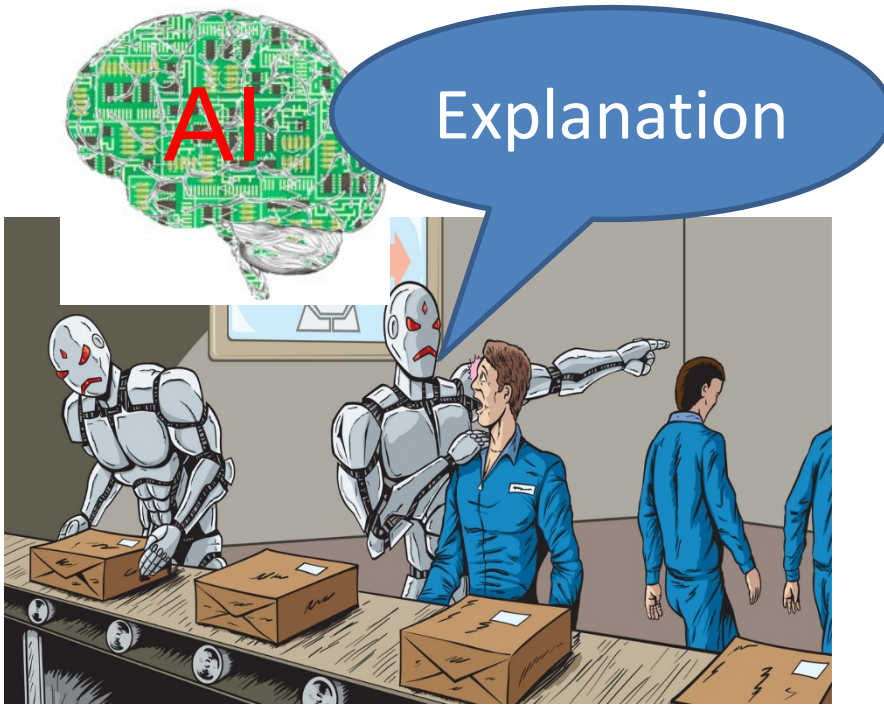
Davenport says we can find new jobs immune to AI invasion, but…

- Jobs of developing AI
  - However, AI itself will develop AI in the near future!



Davenport says we can find new jobs immune to AI invasion, but…

- Jobs of explaining the results generated by AI or action done by AI.
  - Again, however, it will be done by AI!



Hopeless!

How to cope with this situation?

– Basic income ….

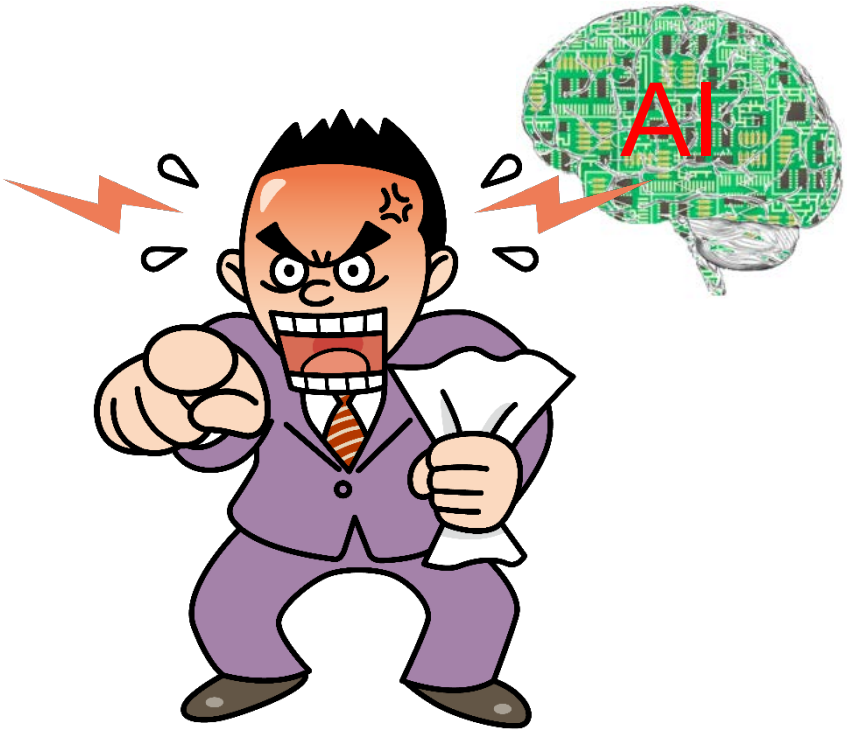
- no incentive,
- Unmotivated…

If a human job is completely replaced by AI and a human being is outside of job process


AI
AI
AI
AI
Shortage of workforce will be saved by AI and Robot.

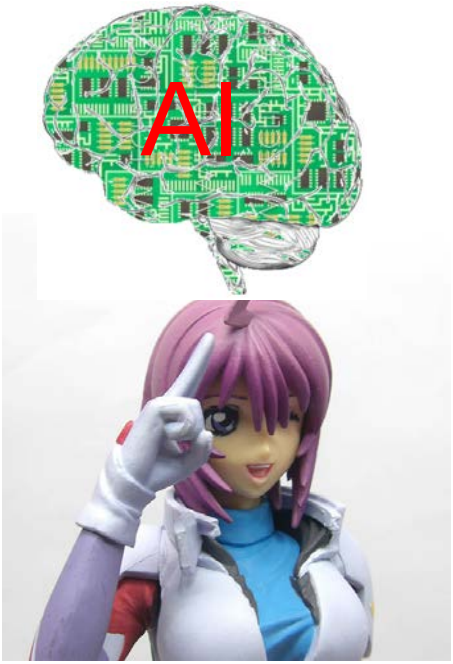
It’s not me but AI says so!

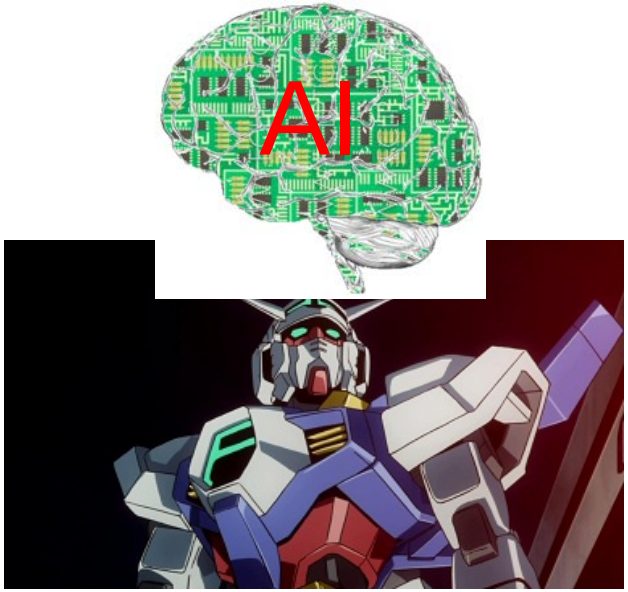
A society without freedom of speech nor even human rights.



➢We need to design the society in which we have the right to object AI’s decision.

Fight against AI with AI





Amplify vs. Replace

- Does Artificial Intelligence amplify human competence?
  - IA: Intelligent Assistance/ Intelligence Amplifier
- Does Artificial Intelligence replace human?
  - AI: Artificial Intelligence
    - "IA vs AI" is a basic view of Markoff's book: Chapter 4 beautifully describes this view and the history of AI.

- IA vs. AI : This is a 60 year old hostile or complementary relation since the starting time of AI.
  - When AI is high, IA is low, or vice versa.
  - If AI technology becomes stuck, researchers become favorite IA.
  - Question: Is Deep Learning IA or AI?

Is a human in the loop or out of the loop?

- Designing principle parallel with “IA vs. AI” is “a human is in the loop or out of the loop .”

- In the loop → IA
  - A system is an extension of human abilities. A human being and a system do work collaboratively.
  - A human does some task being aided by an AI based system.
  - A human might not understand what is going on in an AI system.

- Out of the loop → AI
  - A system acts autonomously. A human only commands a system.
  - A human is no more an actual stakeholder.
  - A human is usually living in a very easy position, but could not get by when some thing happens.

Is a human in the loop or out of the loop?

- In the loop → IA
- Out of the loop → AI

- Generally speaking, the problem is to find the criteria of how far AI system should or can be done on behalf of human beings.
  - Of course, the remaining is to be done by a human being.

- This is a traditional problem between machine and human, but becomes complicated in the era of AI.

- To leave all the decision-making to AI is too easy for a human being but ends up with a kind of addiction and intelligently atrophic(shrinking) .

Example : Self-driving cars

- self-driving: Level 1,2,3,4,5
  - Fully self-driving (no human intervention): Level 4, 5
  - Human intervention is required at a critical moment: Level 3

- If you usually don't drive and are handed over driver's position at the critical moment, can you drive better than AI based self-driving? → NO!

- Thus, Level 3 is much more dangerous than Level 4,5 especially at the moment of taking over, isn't it?

Boundary between amplification and replacement

- It is rather difficult to identify whether a certain IT technology is an amplification or replacement of human ability.

- Examples of apparent amplification:
  - A power suits for an elderly nursing worker.
  - Remote diagnosis by a physician.
  - Augmented reality such as Google glass.

- Examples of apparent replacement
  - Autonomous nursing robot
  - Autonomous door-to-door delivery robot
  - Autonomous scoring and pointing weak point system based on AI

Boundary between amplification and replacement

- Ambiguous example
  - An AI system which is usually working autonomously, but sometimes requests us information: semi interactive system.
  - Example: Question answering system such as tax return assistance
  - Example: Recipe generation system which changes recipes depending on the user's taste and physical condition, and the remained food in the refrigerator
  - self-driving car of Level 3.

Boundary between amplification and replacement

- Considering these examples, one candidate of boundary is the following:
  - If a human user of the system can understand, and predict the system's behavior , then it is an amplification,
  - otherwise , say the system is a completely black box for the user, it is a replacement.
- The examples of previous page
  - An AI based Q/A system can't complete the task by itself, and require the user additional data, information or even instruction.
  - The user thinks he controls the behavior of the AI system, but in reality the system controls the user.
  - While an AI system seems to behave autonomously, it abandons the task at critical moment.
- are somewhere between amplification and replacement.
- Obviously, these example contains essential clue about how AI becomes General AI.

Boundary between amplification and replacement → how to cope with?

- So far, we only think of the solution by technology.
- But if a technological solution is not enough,

➢ we have to think of legal system to prevent the bad consequences.

  - Designing insurance policy to cope with an accident
  - Typically , the problem of self-driving car on the next page.

How to act the critical situation while driving automatically?

◆ When a kid suddenly runs in front of Level 4, say AI driving car,

1. if it avoids collision with the kids, the car jumps into the opposite lane and crash head-on with an approaching car and the person on the car will die,
2. if it goes straight, it collides to the kid and he will be killed.

- What AI driving car should do in this case?
  - AI driving car is a much safer(=lower accident rate) car than a human driving car. Focusing too much on this type of case slows down the development of AI self-driving technologies. “Ridiculous” AI researchers might say.
  - If we accept this AI researchers opinion, we have to work out the way how to treat the case of accident in order to maintain our social structure.

- One solution is an insurance policy for an AI based self-driving car.

- The first problem for this type of insurance is who takes responsibility at an accident.

  - a car manufacturer, a dealer or a car owner?
  - The study is going on.

  - To make an insurance policy, we have to estimate the financial value of person. This kind of topic was almost taboo in AI research, but we could not ward it off anymore.
  - Actually, insurance companies have already done this based on expected total income of the rest of life.

- At collision, does AI estimate both of the kid's expected total income and the driver(not driving, only the car owner)'s expected total income, and make decision based on this estimations?
  - In other words, designing insurance policy does not give us the solution essentially.
  - AI researchers have to be keen to this kind of problem.

➢ Essential solution is to design the transportation system in which this kind of accident situation never happens, such as:

  - The transportation management system in which human driving cars and AI driving cars never share the same road at the same period of time.

  ➢ AI can be a good candidate tool for the above designing.

Another case: The right to be forgotten

- The European Court of Justice rules in 2014 that an Internet search engine must consider requests that the requester wants to remove links to web pages where the information he wants to hide, based on the right to be forgotten.

- More than a million requests of removing links has come to Google.
  - To the best of my knowledge, it seems to be the balance between the right to be forgotten and the right of free speech.

- Google employs 50 to 60 specialist to cope with the requests. Even for Google, it is a heavy burden.
  - The promising way is to employ machine learning technologies and use many cases decided by human experts to build the classification system.
  - This classification system roughly classify the requests, and human experts only look closely the boarder line cases.
  - The difficult thing is to clarify the reason why the request is decided to remove or not. The reason written in natural language sentences is to be clearly articulated to the requester.

AI has already been a black box, so what?

- No matter whether AI may be amplification or replacement ,
  - an AI developer has to take responsibility.
  - But, we have already lapsed into the situation where the stakeholder can't grasp what is going on inside an AI.
  - The stakeholders include :
    - Developers of AI
    - Human agents who offer source data which is used as supervised data of machine learning.
    - People such as working at retail shops or advertising for AI system.
    - Users of AI.

➢ It is necessary now to make a legal system which clearly states who is responsible.

Black box AI harmed financial business

- So called AI trader agent which is regarded as a black box AI , has already been working in stock market etc.. AI traders are connected via internet and the speed is one transaction per less than millisecond.

- Trades are proceeding so fast that human trader can't intervene.
- Under this circumstance, a certain type of trade may trigger the bad chain reaction worldwide where human traders can't intervene in time.

  - In May 2010, some AI trader predicted the default of Greek national bond and took out a lot of selling order. As soon as the order was confirmed, every AI trader felt a kind of crisis and started distress sales, ended up with Dow Jones crashed by 1000 points instantaneously. (Barrat Ch.6 and Ch.8)
  - It took a long time to identity the reason of the crash.

Connected AI systems

- It is regarded as a set of optimal actions of many AI agents connected via internet. The key point is that they have the common language that is a stock price and effects their behaviors decisively.
  - One AI system is not that powerful, but many AI systems communicating each other may make unexpected and uncontrollable situations happen as described.
- We, apparently, need the legal system that clarifies who is responsible in unexpected situation such as above described.
- But, a legal system is necessary and sufficient?

We should not be optimistic

- AI researchers often discuss the ethics of AI from the viewpoint of whether a stand alone AI behaves ethically.

- But, once many AI systems that are network connected start to act autonomously or collaborate, we can't manage nor control them.
  - We have not yet found the way to cope with it. If we install a system to cope with it, there surely are bad guys who try to find loophole for money.
  - That bad guy's action is hard to be detected nor verified.
  - Is it possible to develop another AI system who can detect and verify the bad guy's action?

  - Another concern is : if an AI has the potential to learn the way to understand and estimate the current situation and make a good decision, AI is very near to having self-consciousness?

  - Do AI researchers really do research as consciously considering this concern?

Military and AI

- Using AI in military is inevitable because:
  - AI has already been used in various ways such as reconnaissance drones equipped with AI as well as planning or data processing in military.
  - Basically AI technologies are open to public ending up with being caught up easily by any countries.
  - Different from nuclear weapons, AI technologies are expressed as information. Thus they are easily and rapidly spreading via internet.
  - In other words, terrorists, authoritarian state or enemy country are able to get AI technologies to use for military purposes.
    - In reality, Russia has gotten nuclear technology by using spies (Barrat Ch.14). In today's internet era, AI technologies are extremely easily spreading.
    - The US military think if they are behind in AI technology, they become inferior to other countries in military.

➢ As for AI based weapons, Russel's opinion is important (Nature 28 May 2015, http://www.nature.com/news/robotics-ethics-of-artificial-intelligence-1.17611#/russell)

Military and AI

- Some one says if you embed AI the ethical sense, it's OK, but,
  - in the war time, if our AI robots act ethically, we will be definitely defeated when enemy's robots act unethically. So embedding ethics does not work in the war time.

- A big portion of AI research funds has come from DARPA historically.
  - In the cold war time, DARPA gave a big amount of funds to machine translation research in order to build an NLP system which automatically understands huge volume of Russian documents in short time.
  - A developing of Siri is also funded by DARPA .
  - AI technologies developed by DARPA fund are, of course, transferred to use for military purposes.

Military and AI

- AI technologies developed by DARPA fund are, of course, transferred to use for military purposes.

- But, the real problem is that DARPA might classify some of them to keep superiority of the US.
  - Nevertheless, the technologies are vulnerable to leak.

  - Google that is doing research and development of AI secretly (every body know this fact) hires DARPA high rank officers.

- Based on understanding these things, people including AI researchers, of course, should discuss the dual use of AI.
  - Unfortunately, this kind of discussions are not so well proceeding or sometime even biased.

What happens after AI takes over human jobs?

- If AI takes over human jobs too much, people are expected to be lose their skills or competences.
  - Many persons become unable to travel nor even move in city nearby without their smartphone.
  - Many young workers only wait until the order coming from computer.
  - If a full self-driving car becomes widely used, car driving insurance companies get severer damage.
    - People who have driving skill would rapidly decrease.
    - The accident rate of AI based self-driving car is much lower than human driver.
    - Finally people are not willing to get a driver license.

What happens after AI takes over human jobs?

- If AI takes over human jobs too much, people are expected to be lose their skills or competences.
  - A (not highly skilled) lawyer is facing tough time!
  - They become too much depend on an AI based search engine for judicial precedents.
    - Currently, the similarities used in search engine are calculated by the vocabulary appeared on precedents or process verbal.
    - If AI is able to investigate and analyze the background of cases based on social common sense, no ordinary task are remained.
  - The same situation is highly likely to happen on accountants, tax accountants, workers at research and development section.

Notorious effects on human individual by AI: Filter Bubble

- Facebook uses AI to analyze and utilize its users' profile and information about his/her friends.
- The purpose is to recommend individual users the information that seems to be favored by him/her.
- As the results, it becomes much harder for users to access information that is not inferred as his/her favorite.
- Users are within the bubble made by AI system. The information that AI decides users would not like is filtered out and does not come in the inside of this bubble.
- → so called "filter bubble"

Filter bubble

- The filter bubble help develop the opposite to the original idea if Internet : every one can access every information on the world.

- Since users can't access variety of information, they lose diversity and flexibility of thinking pattern, say informational degeneration.
- It is even harmful to encourage the bias or hate.
  - People sometimes prefer this kind of information degeneration because they don't like to think things sincerely and deeply with much effort.

  - In addition, for IT service provider such as SNS, search engine, it is much efficient and effective advertisement strategy only to show the users' favorite information.

Filter bubble

- However, showing only users' favorite information goes against the people's right to know, and might lead to the domination of information by a few people.

- Some IT company say that every thing is done by software(including AI) and we didn't intervene at all.
- For their viewpoint in terms of law, it is OK for them, but is it really OK for healthy society?

- We need an web tool which gives us the ability to easily access the information we are not necessarily familiar with nor favorite.
- But, is it every thing? The profound cause of filter bubble is in people's psychology, or even human nature. No end in sight….

AI and privacy

- The right to be forgotten.
  - In the case that individual user sends the request of deleting link to his/her individual information , the search engine company should decide whether accept and delete the link or refuse.
  - This decision depends on the balance between the right to be forgotten and the right to know (or how important the search engine is publicly)
  - The decision rule might be worked out by AI (machine learning) technology with huge amount of decision historic data by legal specialists working at the search engine company.
  - Actually, this is a good and useful usage of current AI.

- Profiling
  - Search engines or SNS build up their users' profile with AI technology.
  - User profiles are used in advertisement business which brings them huge amount of profit.
  - At the same time, individual user is very anxious about how their private information will be used.
  - AI creates a bright side and a dark site at the same time.

AI and privacy

- When an appearance of AI robot looks like human and speaks as human being does,
  - a user trusts the robot because a robot does not lie,
  - feels an affinity because of its human like appearance,
- he/she tends to speak his/her own privacy to an AI robot.

- AI robots can collect personal or even privacy information if they want.
  - In order AI not to collect privacy information, the AI is to be equipped with human common sense as to what is a privacy.
  - What is privacy varies person by person, situation by situation.
  - Going to restaurant with a lady is sometimes privacy for the gentleman??. → AI with good common sense! Or general AI!

Misuse of AI

- An AI is easily expected to be used for some wrong or harmful purposes by a bad guy. It is crime and punished by a law.

- Minor but implicit harm is(as stated before):
  - A superior forces the subordinate to do a wrong or impossible task by saying “This is what AI concludes the best or needed.”
  - Maybe the superior does not have a logical reason, but just from his feeling or sentiment. The subordinate gets suffered by a superior’s misuse of AI.
- It can be exaggerate but we need “veto pawer to AI’s decision (human right?!?!).”
- At least, AI has to have a skill to explain the reason why the decision the AI made.

- The issue is two fold:
  - accountability of AI
  - Trust for AI
- Both of them is difficult to implement at this moment.

Self-consciousness and identity

- AI is very near to take an action that is triggered by positive / negative feeling to us human beings. If an AI acts like this, AI becomes General AI in other words “singularity.”

- If an AI gains self-consciousness and identity, it becomes General AI.

  - If an AI is designed to survive, the intension to survive is regarded as its identify. The intension to survive requires to know the state of AI itself. This is self-consciousness.
  - Then, AI has a positive feeling to a person who wants AI to survive, a negative feeling who don’t want AI to survive.
  - As the results, AI acts according to this positive or negative feeling.

- This is a scenario Barrat and Markoff showed in their book.
- Unfortunately, they have not yet indicate the solution to this problem.

Self-identification

- Why , in the morning wake up, are we feel we are the same person as we got asleep yesterday night? → mystery of self-identification
- Kurzweil suggests: A human person gradually replace his/her small part of body and brain, he/she still think he/she is the same person, namely maintains self-identity.
- If this gradual replacing finally makes him/her a post human, he/she(=human2.0) keeps self-identity with no obstacles.
- Thus, a post human being has the same self-identity as current human.→ A post human being will think as we, human being, do. Really???

A tentative summary

- If AI is really needed to do a specific task such as elderly care, we have to investigate more on AI technology useful in that task.

- A new AI technology is needed to explain or visualize for understanding AI's acts. For this, first of all, AI has to observe its internal behavior.
  - i.e. AI should explain its decision of erasing link on the request of erasing the requester.

  - This observation and explanation mechanism is especially important in the field of AI based trading where many AI traders are acting interactively and they are connected via internet with stock prices as their common language.

- We need an AI which protects the problem caused by other AI's
  - Such as alleviating filter bubble effects.

- Of course, we need discussion if we adopt legal constraints.
  - The examples of the same type is a nuclear weapon restriction, international laws for war time, etc. . They have been conducted but there still is a long way to go.

A tentative summary

- Considering the problems of this slide, before saying exaggerate term “singularity”, we have many things to think and try to find the solution.

- Then, after singularity, what will happen?
- A part of answer is written in Nick Bostrom’ s book “Superintelligence.”